DOCUMENT RESUME

ED 103 419 SP 009 026

AUTHOR Trzasko, Joseph A.
TITLE Administration and Research of Competency-/Performance-Based Teacher Education Programs.
PUB DATE [75]
NOTE 30p.
AVAILABLE FROM Dr. Joseph A. Trzasko, Department of Psychology, Mercy College, Dobbs Ferry, New York 10522 (No price quoted)

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS Educational Assessment; *Evaluation; *Performance Based Teacher Education; Post Testing; Pretesting; Program Design; Reliability; *Statistical Analysis; *Teacher Education; Validity
IDENTIFIERS *Mercy College

ABSTRACT

This paper describes the proposed Assessment Center at Mercy College in Dobbs Ferry, New York. The center is intended to provide statistical and technical support for the Mercy College elementary education, special education, and speech and hearing departments in the areas of student assessment, student guidance, and program evaluation. Evaluation and assessment within the Mercy College undergraduate education program is to consist of three phases: preassessment, competency- and performance-based assessment, and postassessment. The paper presents the procedures, instruments, and statistical analyses to be used in each phase. (DDO)

ED103419

ADMINISTRATION AND RESEARCH OF COMPETENCY-/PERFORMANCE-BASED TEACHER EDUCATION PROGRAMS

U.S. DEPARTMENT OF HEALTH EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

BEST COPY AVAILABLE

Joseph A. Trzasko, Ph.D.
Dept. of Psychology
Mercy College

The following issues are explored: (1) general characteristics of competency-/performance- based teacher education programs, (2) application of relevant reliability and validity statistical measures, and (3) application of conventional statistical procedures for student assessment and program evaluation.

In surveying competency-based teacher education (CBTE) programs, the need for a centralized system became evident for within-college program coordination. A case in point is Florida International University's establishment of an Assessment Center:

> The Assessment Center serves as a central data collection/processing location for the School of Education...The large number of students in the various programs, and the volume of enablers to be evaluated, necessitated the Center's existence (Gay, 1973, pg. 11).

The University of Toledo has also adopted a centralized CBTE assessment center based on the following decision:

> It is important to point out now that it is simply not practical to handle the immense amount of data generated by a CBTE program by manual means. Electronic data processing is necessary (Gentry et al., 1974, pg. 1).

Furthermore, Joyce (1974, pg. 95-96) stated:

> It is not possible to imagine competency-based education on a large scale unless it is supported by a computerized management system...In relating 500 to 1,000 teacher candidates to a program containing that many elements requires information access, the coordination of support materials, and the coordination of faculty of very great complexity.

It should also be noted that New York State's Education Department Division of Teacher Education and Certification (1974, pg. 9) requires that "responsibility for information collection has been assigned to a specific person, office, or other organizational unit" be included in CBTE proposals.

Given the above considerations, an Assessment Center at

SP009 026

BEST COPY AVAILABLE

Mercy College is tentatively planned to be operational by September 1, 1975. The Assessment Center is intended to provide statistical-technical support for the Mercy College Elementary Education, Special Education, and Speech-Hearing Departments in the areas of student assessment, student guidance, and program evaluation. Specifically, responsibility of the Assessment Center is to include: (1) updating and publication of general program objectives and operationalized objectives, i.e., specific course objectives; (2) administration and interpretation of pre-assessment instruments to incoming students; (3) scoring and evaluation of competency-based tests, including pre- and post-diagnostic assessment; (4) development and application of competency-based test item statistical procedures for reliability and validity; (5) formation of a bank collection of acceptable and unacceptable competency-based test items; (6) development of assessment procedures for essay and open-ended short-essay questions; (7) development and/or research of teacher behavior taxonomy with regard to performance observation of student teachers; (8) formation of individual student and class group data for use in student guidance; (9) research the relationship between student teacher performance behavior and elementary pupil cognitive and affect changes; (10) administration and interpretation of post-assessment instruments to exiting students; (11) development of follow-up data to supplement program evaluation; (12) publication of research findings; (13) application for educational research grants; and, (14) providing consultation services. Thus, CBTE data collection and interpretation is to be centralized within the Mercy College Assessment Center.

Evaluation and assessment within the Mercy College undergraduate education program is to consist of three phases: (1) pre-assessment, (2) competency- and performance-based assessment, and (3) post-assessment.

Phase I - Pre-Assessment

Pre-assessment of incoming students is based on two

BEST COPY AVAILABLE

assumptions: (1) Since communication is an essential part of teaching, incoming students must demonstrate proficiency in written and spoken English, handwriting, mathematics, and reading. And, (2) exposure to various educational environments is a need of the potential teacher if he is to achieve a better understanding of the range of education and a more realistic approach to the decision of selecting a well-suited program in education. Table 1 outlines the pre-assessment instruments, conditions under which pre-assessment is to be conducted or obtained, and the minimal criteria level required.

(Insert Table 1)

Pre-assessment data is to be obtained for all incoming students, both freshmen and transfer students.

Pre-assessment scores will be stored by the Assessment Center and forwarded to each student's academic advisor for student guidance. Students can also request pre-assessment test scores directly from the Assessment Center.

Research on pre-assessment data includes: (1) calculation of intercorrelations and prediction equations between the various pre-assessment measures; (2) calculation of correlations and prediction equations between the various pre-assessment measures and students' high school average; and, (3) calculation of correlations and prediction equations between the various pre-assessment measures and students' subsequent college grade point average (GPA). Specifically, the Pearsonian product-moment correlation coefficient (r) and related X' and Y' prediction equations (Ferguson, 1971, pg. 96-119).

Phase II - Competency- And Performance-Based Assessment

Competency- and performance-based assessment procedures are the end-products of the following developmental sequence: (a) given the general program objectives, (b) the instructor identifies specific course objectives, (c) which are then operationalized through outlined lecture, reading, and other related materials. As indicated below, the competency- and performance-based assessment procedures directly reflect the stress placed on the course objectives and material. Thus, each assessment

BEST COPY AVAILABLE

TABLE 1. Pre-Assessment Instruments.

Pre-Assessment Instruments	Conditions Under Which Pre-Assessment Instruments Conducted Or Obtained	Minimal Criteria Level
Scholastic Aptitude Test (SAT) Verbal & Quantitative Scores	SAT scores obtained through Admissions Office or scores provided by Educational Testing Service through student's request.	No criteria beyond Admission Office's standards. SAT scores to be used as student characteristic data material.
College Student Questionnaire (CSQ)	CSQ will be administered and interpreted through Assessment Center on publicized dates.	No criteria used. CSQ scores to be used as student characteristic data.
Spoken English Proficiency Test	Screening will be performed by Mercy College Speech-Hearing Department. The speech pattern of the student seeking certification as a teacher shall exhibit the following qualities: (a) pleasantness (satisfying to the listener in the areas of voice and diction); (b) efficiency (transmits the message intelligibly); (c) freedom from peculiarities of voice or diction that identify the speaker ethnically or geographically.	Each student shall read a given selection to a panel of three members of the Speech-Hearing Department for judgment in the following areas: (1) Voice (pitch, rate, & volume) judged as adequate or inadequate for classroom performance; and (2) Diction (consonant production, lisping, dentalization, voicing, omission, vowel production, substitution, dipthongizing, and nasalizing) judged as acceptable or unacceptable based on the absence of error, occasional error, or persistent error. Persistent error (repetition of an error 30% of the time in any category) shall render the pattern unacceptable and in need of remediation.

BEST COPY AVAILABLE

TABLE 1. Continued.

Pre-Assessment Instruments	Conditions Under Which Pre-Assessment Instruments Conducted or Obtained	Minimal Criteria Level
Written English, Reading, and Mathematics Proficiency Tests - California Achievement Tests (CAT) Level 5	CAT will be administered and interpreted through Assessment Center on publicized dates.	Incoming students must demonstrate at least 30th percentile on each test with regard to grade 12 norms. Students falling below the 30th percentile on any test will be required to attend remedial courses or pursue independent study on the identified deficiency area. These students will be retested on the identified deficiency area by the end of their first academic year.
Handwriting Proficiency Test	Student is to handwrite a given printed passage at the normal rate. The printed passage is devoid of grammatical punctuation. The Handwriting Proficiency Test will be administered and interpreted through Assessment Center on publicized dates.	Legibility is to be judged by three members of the Education Department as acceptable or unacceptable for classroom. Minimal criteria is an acceptable rating by 2 of 3 judges. The number of grammatical punctuation errors is not to exceed 30%.
Exploration of at least one of the following five situations: (1) early childhood education, (2) urban education, (3) open education, (4) bilingual education, and (5) special education (including mentally handicapped, emotionally handicapped, or speech & hearing handicapped)	Arrangements for observations of an education situation either made independently by student or through Mercy College Education Department.	Acceptable completion of an observation record regarding facilities, personnel, class structure, classroom management routines, etc.

BEST COPY AVAILABLE

procedure provides a diagnostic feature for student guidance. The following competency- and performance-based assessment procedures were developed from guidelines provided by the Multi-State Consortium on Performance-Based Teacher Education (1974) publication on assessment and the Gentry et al. (1974) description of the University of Toledo's CBTE program.

The assessment procedures reflect both the preference of the instructor and course objectives. Assessment procedures are of three major types: (1) selected response, i.e., multiple-choice, fill-in, etc.; (2) essay response; and, (3) performance response, i.e., observation checklists, demonstration-project rating scales, etc. In consultation with the faculty of the Mercy College Education Department, courses were described according to the preferred student assessment instruments. While some courses rely on only one assessment procedure, other courses utilize a combination of the assessment instruments. For example, Ed. 32 Psychology of Learning relies on selected response, while Ed. 40 Reading Methods relies on a combination of selected, essay, and performance responses.

The following describes each assessment procedure and the appropriate statistical-experimental design.

Selected Response

Courses in which the instructor wholly or partially relies on short-answer examinations utilize, e.g., a 70% criteria level on competency-based post-examinations. The general experimental design is presented in Figure 1. The specific course

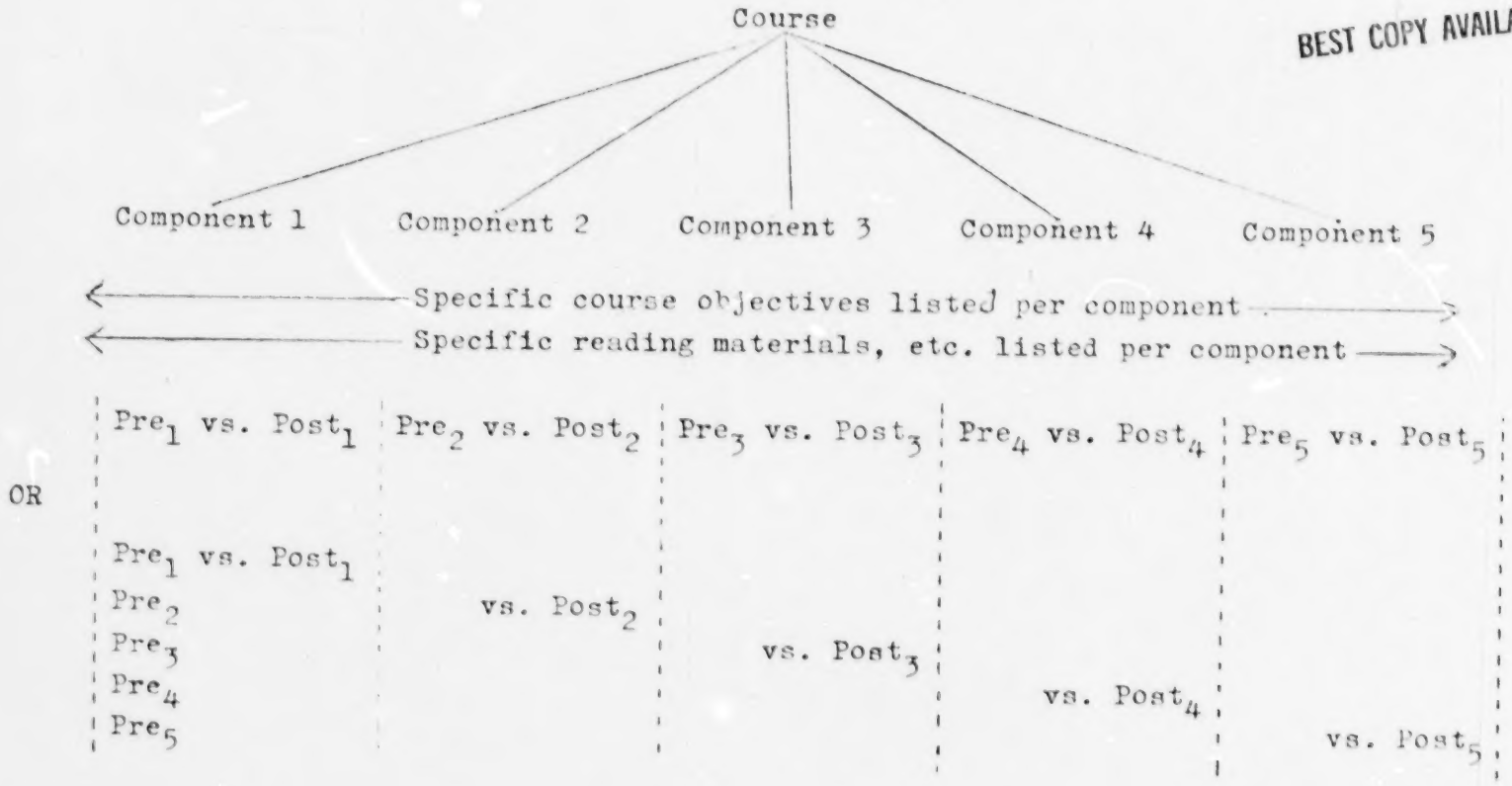
(Insert Figure 1)

outline for Ed. 32 Psychology of Learning is presented in Figure 2.

(Insert Figure 2)

The general program competency for Ed. 32 Psychology of Learning requires "knowledge of learning theories." This general competency can be operationalized into the following four course objectives: (1) information of historical background, experimental procedures, and comparison of classical and operant conditioning within laboratory setting; (2) specific information on experimental learning procedures, including parameters of reinforcement,

FIGURE 1. Module components of a course including pre- and post-examinations.

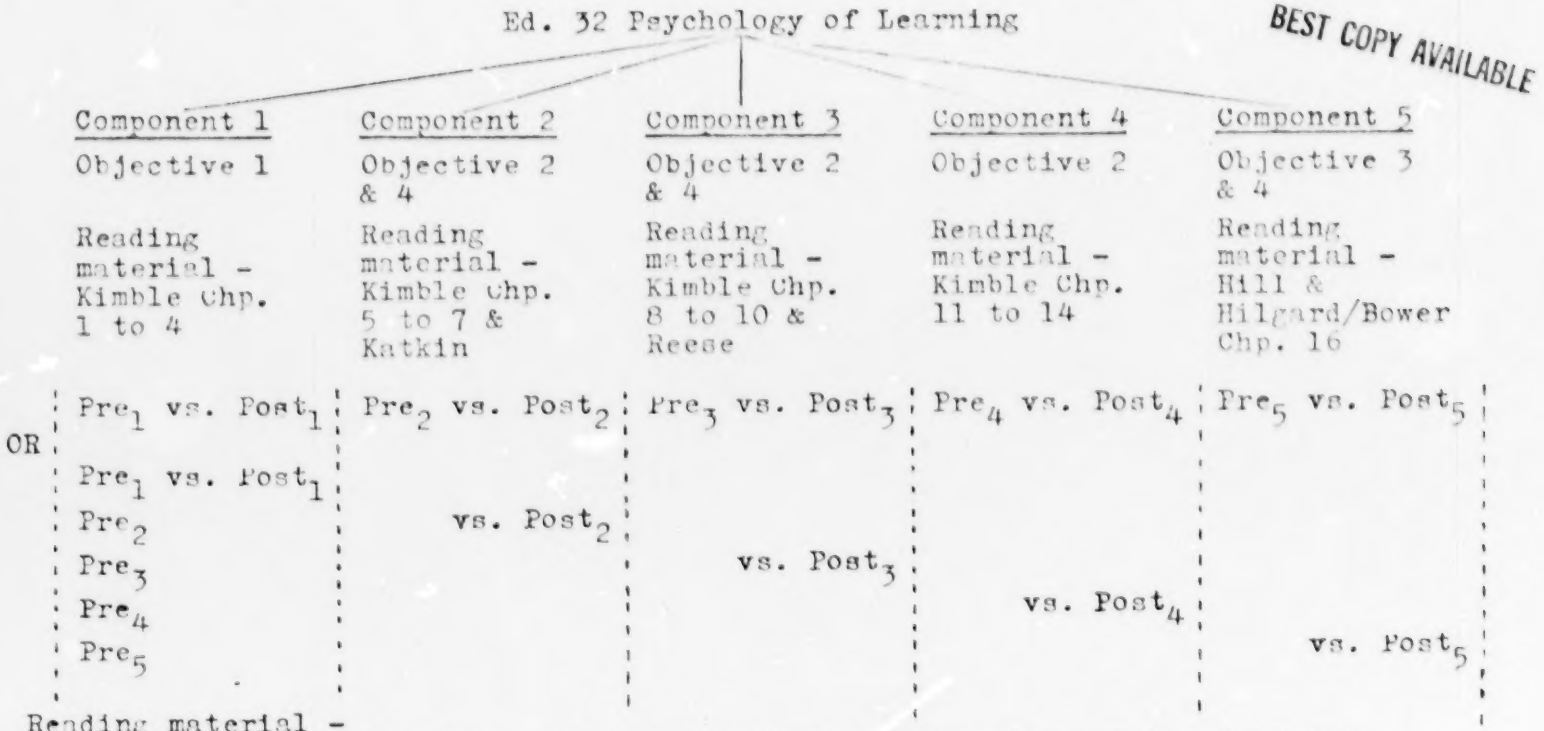


BEST COPY AVAILABLE

Student given credit for the course when successfully completes all components within a specified time limit with, e.g., a 70% criteria level per each post-examination or a 70% average across post-examinations.

FIGURE 2. Example of module components of a course, Ed. 32 Psychology of Learning.

BEST COPY AVAILABLE


Ed. 32 Psychology of Learning
Component 1
Objective 1
Reading material - Kimble Chp. 1 to 4
Component 2
Objective 2 & 4
Reading material - Kimble Chp. 5 to 7 & Katkin
Component 3
Objective 2 & 4
Reading material - Kimble Chp. 8 to 10 & Reese
Component 4
Objective 2
Reading material - Kimble Chp. 11 to 14
Component 5
Objective 3 & 4
Reading material - Hill & Hilgard/Bower Chp. 16
Pre1 vs. Post1
OR
Pre1 vs. Post1
Pre2 vs. Post2
Pre3 vs. Post3
Pre4 vs. Post4
Pre5 vs. Post5

Reading material -
Kimble, G.A. Hilgard and Marquis' conditioning and learing. N.Y.: Appleton-Century-Crofts, 1961.
Katkin, E.S. Instrumental autonomic conditioning. N.Y.: General Learning Press, 1971.
Reese, E.P. The analysis of human operant behavior. Dubuque, Iowa: Wm. C. Brown Co., 1966.
Hill, W.F. Learning: A survey of psychological interpretations. Scranton, Pa.: Chandler Publishing Co., 1971.
Hilgard, E.R. & Bower, G.H. Theories of learning. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1975.

BEST COPY AVAILABLE

secondary reinforcement, general theories of reinforcement, mechanisms of reward, nature of extinction, generalization, discrimination, motivation, and application to personality; (3) information on general learning theories, including Pavlov, Skinner, Hull, Guthrie, Tolman, Mowrer, Estes, etc.; and, (4) application of learning theory to educational, medical, and clinical situations.

For each course component, competency-based pre- and post-examinations are to be based on and reflect content stress, e.g., for component 1: approximately 15% of lecture and examination based on Chp. 1 The Definition of Learning, approximately 25% of lecture and examination based on Chp. 2 Conditioning in Historical Perspective, approximately 30% of lecture and examination based on Chp. 3 Classical and Instrumental Conditioning Experiments, and approximately 30% of lecture and examination based on Chp. 4 Classical and Instrumental Conditioning Compared.

The competency-based pre- and post-examinations should be as objective as possible, e.g., multiple-choice, fill-in, and short-answer. True and false questions are to be avoided due to the high guess factor. Multiple-choice questions are preferred due to possible machine scoring and application of correction for guessing formulas (Downie, 1967, pg. 145-146).

Students falling below the 70% minimum criteria level can be recycled back through the specific diagnosed area(s) of difficulty. A major difficulty with recycling students is the need for several alternative test forms; otherwise, the student may adopt a rote memory strategy of the correct answers. If this should occur due to the use of only one test form, the test then becomes a measure of the student's rote memory ability rather than of knowledge competency. This difficulty of a student adopting a rote memory strategy is applicable to using the same test form in both the pre- and post-examinations. However, a 'cross-over design' may be used (Wike, 1971, pg. 64). Specifically, half of the class receives test form 1 and the other half of the class receives test form 2

BEST COPY AVAILABLE

at pre-examination; each half of the class then receives the other test form at post-examination.

The competency-based pre- and post-examinations are to be administered within the class by the instructor. A copy of each competency-based examination is then to be filed with the Assessment Center and the student answer sheets are to be machine scored when possible. Results from each competency-based examination are then to be stored by the Assessment Center and also forwarded to the instructor for student guidance.

Therefore, pre-, post-, and change scores are available per course component per student for each course. A pre- vs. post- correlated t-test (Ferguson, 1971, pg. 153-155) or Walsh nonparametric (Siegel, 1956, pg. 83-87) comparison is then possible per component. It is therefore statistically possible to conclude whether or not a significant increase in amount of student knowledge has occurred. This is illustrated in Figure 3.

(Insert Figure 3)

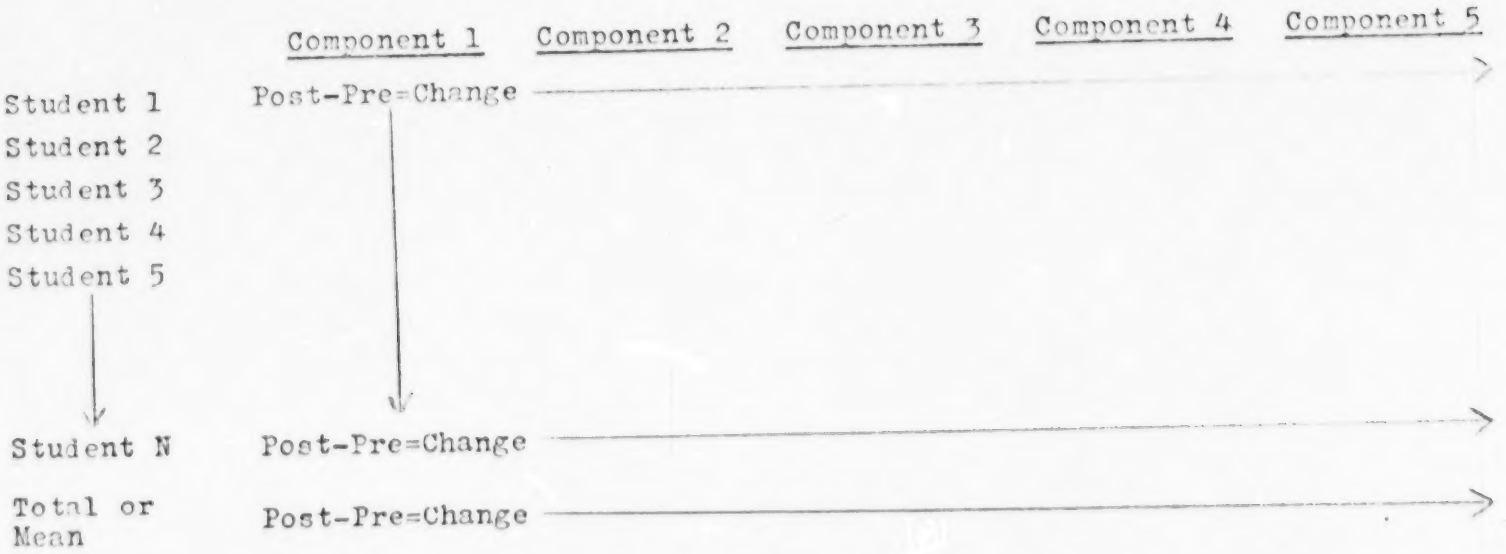
Diagnostic analysis of pre- competency-based test results, e.g., percent incorrect per test area, is of value since it is then possible for the instructor to direct course stress toward identified area(s) of difficulty for the class in general.

Validity analyses of each test item can be obtained by the following three measures: (1) difficulty and discriminating power indicies (Gronlund, 1965); (2) point-biserial correlation (r_{pb}), i.e., the relationship between dichotomous correct-incorrect item scores vs. the continuous total test score; and, (3) phi coefficient (r_{ϕ}), i.e., the relationship between dichotomous correct-incorrect item scores vs. the dichotomous pass-fail test scores. A more complete treatment of the point-biserial correlation and the phi coefficient can be found in Cornell (1956, pg. 307-312), Edwards (1967, pg. 122-128), Ferguson (1971, pg. 356-358), and McNemar (1962, pg. 188-198).

Reliability analysis of each test can be obtained

BEST COPY AVAILABLE

FIGURE 3. Experimental design statistically comparing pre- vs. post- competency-based examinations.



Pre-, post-, and change scores obtained per student per course component. Pre- and post- scores to be supplied either by instructor or scored by Assessment Center.

Statistical analyses - To indicate effectiveness of course material and/or instructor: (a) pre- vs. post- correlated t-test or Walsh nonparametric comparison per component; and, (b) change score vs. zero change t-test comparison per component.

Statistical conclusions - To indicate whether or not significant increase in knowledge acquired by student, i.e., course and instructor effectiveness.

BEST COPY AVAILABLE

from the Kuder-Richardson formula 20, which applies to dichotomously scored items. Furthermore, the Kuder-Richardson formula 20 results are compatible with the split-half method of estimating reliability (Ferguson, 1971, pg. 265-368).

It should be noted that the above validity and reliability analyses are employed within the University of Toledo CBTE program (Gentry et al., 1974).

Essay Response

Courses in which the instructor wholly or partially relies on essay responses utilize scales to reflect student performance on the specified criteria. That is, assessment of essay examinations involves a scale rating of the primary criteria indicated by the instructor. Figure 4 is an example of a basic essay evaluation scale.

(Insert Figure 4)

The instructor indicates the minimal acceptable criteria level per scale, e.g., 70% or 7 point scale score. The diagnostic feature allows identification of a student's difficulty as a deficiency in either level of knowledge, interpretation, and/or application.

An alternative assessment approach is presented in Figure 5. Short-essay responses to open-ended questions can be evaluated according to a scale indicating the student's level of comprehension. The instructor rates each answer

(Insert Figure 5)

as to the student's displayed level of comprehension. Questions may include, e.g., definition of concepts, relationship between concepts, etc. The instructor indicates the minimal acceptable criteria level per scale, e.g., 70% or 7 point scale score. The diagnostic feature allows identification of the specific concepts, etc. with which the student has displayed difficulty.

The instructor forwards completed essay assessment scales to the Assessment Center, where they are stored and a summary record is constructed per student per course.

FIGURE 4. Example of an Essay Evaluation Scale.

ESSAY EVALUATION SCALE

BEST COPY AVAILABLE

(1) Amount of Relevant Knowledge Facts (Level of Knowledge Facts)

0	1	2	3	4	5	6	7	8	9	10
no information relayed or information totally not relevant				information incomplete and partially relevant					information complete and relevant	

(2) Interpretation or Explanation of Subject Matter (Level of Interpretation)

0	1	2	3	4	5	6	7	8	9	10
no interpretation or totally inadequate interpretation				interpretation partially adequate and relevant					interpretation completely adequate and relevant	

(3) Application of Subject Matter (Level of Application)

0	1	2	3	4	5	6	7	8	9	10
no application attempted or totally inadequate application				application partially adequate and relevant					application completely adequate and relevant	

FIGURE 5. Example of Open-Ended Short-Essay Scale.

OPEN-ENDED SHORT-ESSAY SCALE

Ques. 1

0	1	2	3	4	5	6	7	8	9	10
no comprehension				partial comprehension					complete comprehension	

(Level of Comprehension)

Ques. 2 etc.

BEST COPY AVAILABLE

Performance Response

Courses in which the instructor wholly or partially relies on performance responses utilize scales to reflect student performance on specified criteria. That is, evaluation of student performance involves a scale rating of the primary criteria indicated by the instructor. Performance responses are to be applied in the three following cases: (1) demonstration and project performance, (2) laboratory observation performance, and (3) student teaching performance.

Demonstration and Project Performance

The instructor specifies both the relevant scales in addition to the minimal acceptable performance level. For example, Figure 6 presents the criteria used to measure the degree of competency in oral performance for D 60 Oral Performance of Literature. Students are required to present

(Insert Figure 6)

six to eight different problem readings from prose to poetry, which is to be prepared outside of class for within-class presentation. For each rating scale, a score of two is the minimal acceptable performance level.

Spe 55 Voice and Speech Science requires written report projects on, for example, the evolutionary development of speech mechanisms, singing vs. speech voice, or voice and sound in the communication process of other mammals. A 0 - 10 rating scale with a minimal scale score of seven can be applied with regard to the following criteria: (1) relevant background or previous research findings and theories, (2) logical interrelationship among the previous findings, (3) formation of general trends or hypotheses, and (4) adequate research writing style.

The instructor forwards completed demonstration and project evaluation scales to the Assessment Center, where they are stored and a summary record is constructed per student per course. Furthermore, interrater or scoring reliability is available by obtaining simultaneous or independent ratings from other judges. Possible statistical analyses include: (1) repeated measurements analysis of variance to indicate whether

FIGURE 6. Example of a Demonstration and Project Assessment Rating Scale for a course, D 60 Oral Performance of Literature.

DEMONSTRATION AND PROJECT ASSESSMENT RATING SCALE

	0 Totally Unacceptable	0.5	1.0 Below Average	1.5	2.0 Average	2.5	3.0 Above Average	3.5	4.0 Excellent
I. Physical Performance									
1. Body tension (relaxation, as well as absence of self-consciousness is ideal)						etc.			
2. Posture (comfortable stance, yet strong and supportive)									
3. Energy (alertness and enthusiasm included)									
4. Audience contact (awareness of audience and free from dependance on text/script)									
II. Oral Performance									
1. Vocal variety									
(a) Pitch									
(b) Rate									
(c) Volume									
2. Vocal quality									
(a) Sound Production									
(b) Resonance									
(c) Breathing									
3. Projection									
4. Articulation (clarity; appropriate accent/ dialect)									
III. Interpretative Performance									
1. Theme (accuracy and clarity)									
2. Mood/Atmosphere (accuracy and clarity)									
3. Characterizations (distinct, individual, appropriate)									
4. Emotion (believability, honesty)									
5. Builds (clarity and appropriateness)									
6. Transitions (clarity and accuracy)									

BEST COPY AVAILABLE

BEST COPY AVAILABLE

or not student performance per scale significantly changes over the different projects, e.g., readings (Ferguson, 1971, pg. 241-243); (2) reliability of raters (interjudge agreement) via subject by rater analysis of variance (Downie & Heath, 1970, pg. 225-227); and, (3) reliability of a single rater and group of raters (Downie & Heath, 1970, pg. 227).

Instructors also have the option of utilizing a Yes // No // scale regarding successful participation within or completion of the demonstration or project. However, the more extensive rating scales are recommended due to the greater diagnostic capability.

Laboratory Observation Performance

The instructor specifies both the relevant scales in addition to the minimal acceptable performance level. For example, Figure 7 presents the criteria used to measure the degree of competency in the laboratory performance for Ed 41 Reading Laboratory Practicum, which consists of the diagnostic testing and teaching of individual or small groups in the elementary school.

(Insert Figure 7)

Student teachers will be rated on the Laboratory Observation Scale at the beginning and at regular intervals throughout the laboratory experience. The minimal acceptable level per scale is: a score of 3 is required for scales 1, 2, 3b, and 6a; a score of 7 is required for scales 4, 6b, and 6c; a score of 70%-30% is required for scale 3a; and, no criteria is required for scales 5a and 5b. The diagnostic feature of the Laboratory Observation Scale allows identification of each student's area(s) of difficulty with subsequent remediation.

The instructor forwards completed Laboratory Observation Scales to the Assessment Center, where they are stored and a summary record is constructed per student per course. Furthermore, interrater or scoring reliability is available by obtaining simultaneous ratings from independent judges. Possible statistical analyses include: (1) repeated measurements analysis of variance to indicate whether or not

FIGURE 7. Laboratory Observation Scale.

LABORATORY OBSERVATION SCALE

BEST COPY AVAILABLE

(1) Selection of Teaching Approach

(1)	(2)	(3)	(4)
Incorrect approach inappropriately applied	Incorrect approach appropriately applied	Correct approach inappropriately applied	Correct approach appropriately applied

(2) Development of Classroom Procedures (Variety of Procedures Used with Same Pupils)

(1)	(2)	(3)
Uses one procedure continuously	Change of procedure without regard to learning situation (child-environment factors)	Changes procedure for maximum effectiveness

(3) Effective Presentation and Controls

(a) Clarity of Presentation

0	10	20	30	40	50	60	70	80	90	100	Logical
100	90	80	70	60	50	40	30	20	10	0	Illogical

Percent of Total Time

Illogical sequence of directions all of the time	Illogical sequence of directions most of the time	Logical sequence of directions most of the time	Logical sequence of directions all of the time

(b) Specificity of Directions

0	1	2	3	4	5	6	7	8	9	10

Unclear all of the time	Unclear most of the time	Clear most of the time	Clear all of the time

FIGURE 7. Continued.

BEST COPY AVAILABLE

(4) Use of Available Resources

0 1 2 3 4 5 6 7 8 9 10

Minimal Intermediate Maximal

(5) Use of Appropriate Questioning Techniques

(a) Indicate the Most Frequent Type of Question Used

// Memory Recall (e.g., definition or information recall)
// Application (e.g., apply previous information to new situation)
// Analysis (e.g., ability to distinguish and compare various concepts)
// Synthesis (e.g., ability to rearrange component ideas into new wholes)
// Evaluation (e.g., ability to make judgments based on internal evidence or external criteria)

(b) Indicate the Least Frequent Type of Question Used

// Memory Recall (e.g., definition or information recall)
// Application (e.g., apply previous information to new situation)
// Analysis (e.g., ability to distinguish and compare various concepts)
// Synthesis (e.g., ability to rearrange component ideas into new wholes)
// Evaluation (e.g., ability to make judgments based on internal evidence or external criteria)

(6) Determination of Diagnostic, Evaluation, and Prescriptive Methods

(a) Determination of Diagnostic Tests

0 1 2 3 4 5 6 7 8 9 10

Student is unable to construct or select appropriate tests

Student is able to construct or select appropriate tests with assistance

Student is able to construct or select appropriate tests without assistance

FIGURE 7. Continued.

(b) Determination of Test Evaluation

BEST COPY AVAILABLE

0	1	2	3	4	5	6	7	8	9	10
Student is unable to interpret diagnostic tests				Student is able to interpret diagnostic tests with assistance					Student is able to interpret diagnostic tests without assistance	

(c) Determination of Prescriptive Methods

0	1	2	3	4	5	6	7	8	9	10
Student is unable to determine prescriptive metnods				Student is able to determine prescriptive methods with assistance					Student is able to determine prescriptive methods without assistance	

BEST COPY AVAILABLE

or not student performance per scale significantly changes over the observation intervals (Ferguson, 1971, pg. 241-243); (2) reliability of raters (interjudge agreement) via subject by rater analysis of variance (Downie & Heath, 1970, pg. 225-227); and, (3) reliability of a single rater and group of raters (Downie & Heath, 1970, pg. 227).

A fundamental assumption of CBTE programs is 'accountability,' i.e., evaluation of teacher competence in terms of student learning (Airasian, 1974). Therefore, pre- and post- pupil performance scores are to be obtained during the laboratory experience to indicate possible cognitive effects of student teacher performance. Pre- and post- pupil performance scores are to be obtained from standardized achievement tests, e.g., the Metropolitan Reading Test (Grades 2-9) for Ed 41 Reading Laboratory Practicum. The instructor is to forward pre- and post- pupil performance scores to the Assessment Center, where they are to be stored and a summary record constructed per student per course. In addition, the Assessment Center is to obtain new and revised standardized tests for use in measuring pupil performance.

Possible statistical analyses include: (1) correlated t-test (Ferguson, 1971, pg. 153-155) or Walsh non-parametric (Siegel, 1956, pg. 83-87) to determine whether or not a significant change has occurred in pupil cognitive levels; (2) Pearsonian product-moment correlations (Ferguson, 1971, pg. 96-106) between observed student teacher laboratory performance and pupil cognitive change; and, (3) coefficients of determination (r^2) and nondetermination ($1 - r^2$) indicating the percent of variation in pupil performance change scores accounted for by student teacher performance (Ferguson, 1971, pg. 115-117). It should be noted that any change or lack of change in pupil performance scores may not solely be attributed to student teacher performance due to the vast variety of other uncontrolled variables. However, the coefficient of determination (r^2) will provide an indication of the contri-

BEST COPY AVAILABLE

bution of student teacher performance.

Student Teaching Observation Performance

The department specifies both the relevant scales in addition to the minimal acceptable student teacher performance on specified criteria. That is, evaluation of student performance involves a scale rating of the primary criteria as indicated in Figure 8.

(Insert Figure 8)

For example, in Ed 60 Supervised Student Teaching, the student teacher will be observed following each four week off-campus experience. The minimal criteria level is 70% or a 7 point score per scale. The diagnostic feature of the Student Teaching Observation Scale allows identification of each student's area(s) of difficulty. The Assessment Center is to maintain each student's rating per scale over observation periods as indicated in Figure 9. Furthermore, each

(Insert Figure 9)

student's performance per scale characteristic can be graphed over the four observation periods as indicated in Figure 10.

(Insert Figure 10)

Scoring alternatives include requiring minimal performance on a given number of scales and/or differential weighting of each scale. Furthermore, interrater or scoring reliability is available by obtaining simultaneous ratings from independent judges. Possible statistical analyses include: (1) repeated measurements analysis of variance to indicate whether or not student teacher performance per scale significantly changes over the four observation intervals (Ferguson, 1971, pg. 241-243); (2) reliability of raters (interjudge agreement) via subject by rater analysis of variance (Downie & Heath, 1970, pg. 225-227); and, (3) reliability of a single rater and group of raters (Downie and Heath, 1970, pg. 227).

Referring to the fundamental assumption of 'accountability,' pre- and post- pupil cognitive (performance) and affect (course and instructor attitude) scores are to be obtained to indicate possible effects of student teacher

BEST COPY AVAILABLE

FIGURE 8. Example of Student Teacher Observation Scales.

STUDENT TEACHER OBSERVATION SCALES

Organizational Skills

(1) Writes objectives appropriate to content area.

(2) Creates systems and schedules for use of materials.

(3) Provides smooth transition between activities.

(4) Works with teacher to develop immediate and long-range plans for the total class.

0 1 2 3 4 5 6 7 8 9 10

Student does not display skill at all

Student partially displays skill

Student competently displays skill all of the time

etc.

Presenting Skills

(1) Conveys ideas clearly and effectively.

(2) Writes legibly.

(3) Varies speech rate and volume.

(4) Integrates musical, artistic, and poetic experiences with cognitive learning.

(5) Uses audio-visual materials to stimulate interest.

(6) Organizes objectives so as to provide a logical order of presentation.

(7) Uses the inductive and deductive methods of teaching.

Responding Skills

(1) Isolates, discusses and solves possible problem areas early in the instruction.

(2) Makes consensus decisions in areas that may be controversial.

(3) Exhibits ability to work positively with individuals from different cultural, social, economic, and ethnic backgrounds in achieving the goals.

(4) Anticipates reactions of pupils.

(5) Plans physical facilities and activities consistent with the needs of the pupils.

Selecting Skills

(1) Identifies skills basic to content area.

FIGURE 8. Continued

BEST COPY AVAILABLE

(2) Chooses educational textbooks and materials appropriate or related to instructional objectives.

(3) Uses verbal and non-verbal cues in the classroom.

(4) Uses community resources as an avenue for learning.

0 1 2 3 4 5 6 7 8 9 10

Student does not display skill at all

Student partially displays skill

Student competently displays skill all of the time

etc.

Developing Skills

(1) Develops alternative plans on short notice.

(2) Designs learning experiences which develop inquiry, decision making, and problem solving.

(3) Demonstrates understanding of content process and materials appropriate to the lesson.

Eliciting Skills

(1) Helps pupils formulate objectives.

(2) Poses a problem to introduce an activity.

(3) Motivates pupils to express their feelings, perceptions, and emotions.

Nurturing Skills

(1) Encourages pupils to share their interests with peers.

(2) Motivates pupils to assume responsibility for self-improvement.

(3) Supports pupils in their learning efforts.

(4) Functions as a facilitator rather than a controller - a helper rather than a director.

Appraising Skills

(1) Constructs pre- and post-diagnostic tests.

(2) Uses a variety of evaluative techniques to assess all aspects of a pupil's learning.

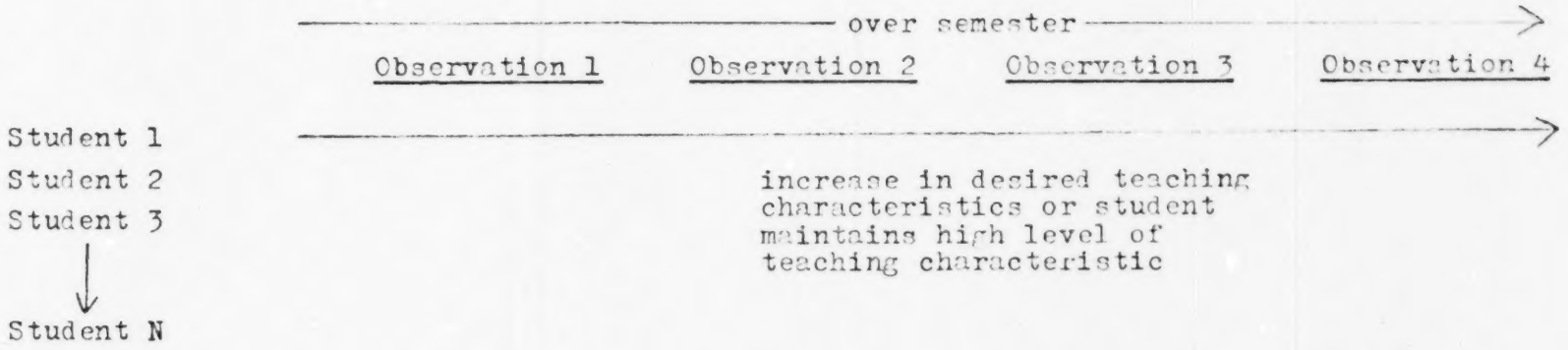
(3) Administers and interprets teacher-made and standardized tests.

(4) Diagnoses pupils' verbal and written responses for future planning.

(5) Prescribes in accordance with test results.

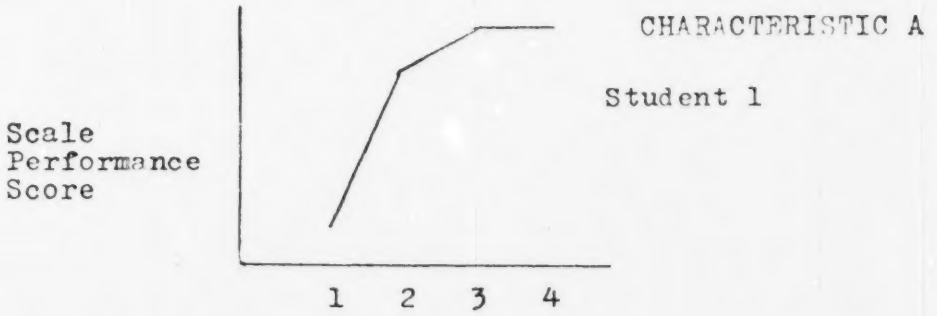
FIGURE 9. Evaluation of Student Teacher Observed Behavior.

BEST COPY AVAILABLE



Statistical analyses - Repeated measurements analysis of variance per each teaching characteristic or total observational score.

FIGURE 10. Graph Per Student Per Characteristic Over 4 Observation Periods.



Statistical conclusions - To indicate change in teaching behavior over student teaching experience.

BEST COPY AVAILABLE

performance. This is presented in Figure 11. The student

(Insert Figure 11)

teacher is to construct and administer pre- and post-competency-based examinations to determine change in pupil cognitive performance.

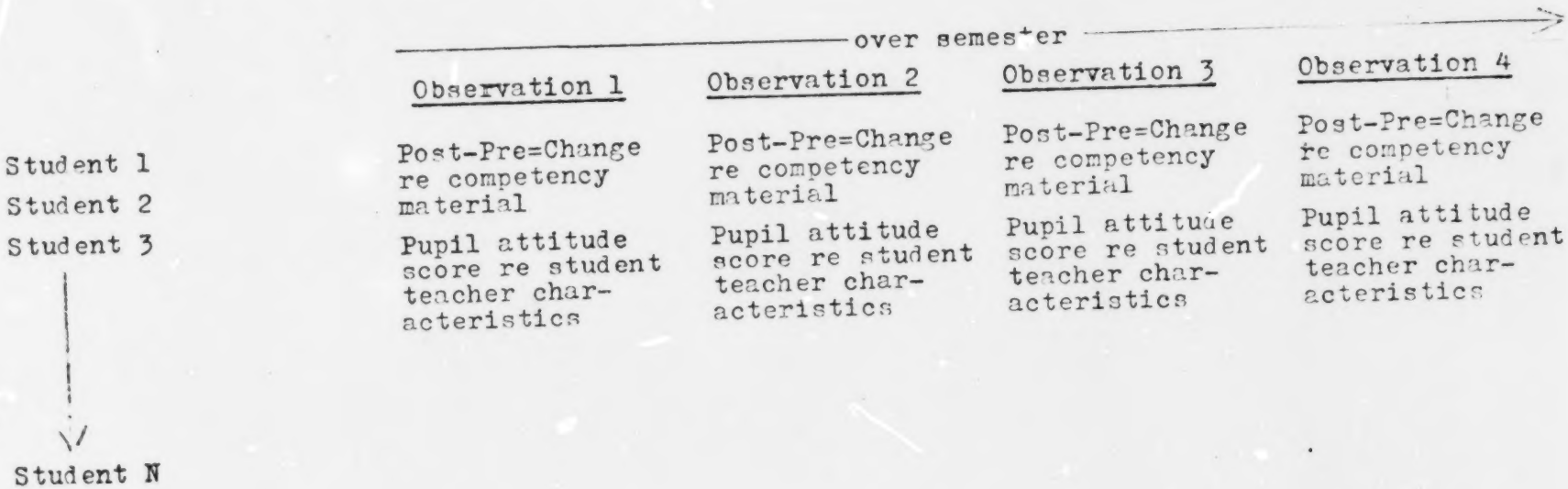
Pre- and post- pupil cognitive and attitude scores are to be stored within the Assessment Center and a summary record is constructed per student teacher. Possible statistical analyses include: (1) correlated t-test (Ferguson, 1971, pg. 153-155) or Walsh nonparametric (Siegel, 1956, pg. 83-87) to determine whether or not a significant change has occurred in pupil cognitive level per observation period; (2) Pearsonian product-moment correlations (Ferguson, 1971, pg. 96-106) between observed student teacher performance and pupil cognitive and affect changes; and, (3) coefficients of determination (r^2) and nondetermination ($1 - r^2$) indicating the percent of variation in pupil cognitive and affect scores accounted for by student teacher performance (Ferguson, 1971, pg. 115-117). As previously indicated, any change or lack of change in pupil cognitive or affect may not solely be attributed to student teacher performance due to the vast variety of other uncontrolled variables. However, the coefficient of determination (r^2) will provide an indication of the contribution of student teacher performance to both pupil cognitive and affect levels. Due to the difficulty in controlling the numerous variables operating upon pupil cognitive and affect levels, future research should be directed toward multi-variate rather than uni-variate approaches.

Phase III - Post-Assessment

In addition to college and department course and credit requirements, exiting students are required to submit scores from the Common and Area National Teacher Examinations (NTE) taken during their last semester prior to graduation. Both the Common and Area NTE are constructed by Educational Testing Service of Princeton, N.J. and provide measures of academic

BEST COPY AVAILABLE

FIGURE 11. Experimental design statistically comparing elementary pupil cognitive gain and affect change regarding student teacher performance.



Pre- and post-test scores obtained from competency-based examinations prepared by student teacher regarding module covered. Since student teachers' effectiveness is expected to increase over the semester, an increase in pupil cognitive change score and an increase in positive pupil attitude toward the student teacher is expected over observation periods. The possibility of confounding cannot be overlooked. For example, confounding may result from the different material covered in each period, possibly different classes in each period, and variables originating from outside of the class.

Statistical analyses - Pre- vs. post- correlated t-test or Walsh nonparametric comparison regarding competency material covered per observation period or change score vs. zero change t-test comparison per observation period.

Statistical conclusions - To statistically indicate pupil cognitive and affect changes over the student teaching experience.

BEST COPY AVAILABLE

preparation.

The Common Examinations provide a general appraisal of the student teacher's professional preparation and general academic attainment. The Common NTE consists of Professional Education Tests (Psychological Foundations of Education; Societal Foundations of Education; and, Teaching Principles and Practices) and General Education Tests (Written English Expression; Social Studies, Literature, and Fine Arts; and, Science and Mathematics). The Area Examinations measure knowledge of the subject in which the student has concentrated and which the candidate intends to teach; therefore, Mercy College Education Department Seniors are to take either Area 01 Education in the Elementary School, Area 02 Early Childhood Education, or Area 32 Education of the Mentally Retarded. Area Examination questions also assess knowledge of instructional methods pertinent to the area of specialization as well as familiarity with major professional organizations and journals.

Since the Common and Area NTE results are to be used for program evaluation, no specific criteria level is required by the student. NTE results per student are to be stored by the Assessment Center and summary analyses prepared. Although the Area NTE provides only a total score, the Common NTE provides a diagnostic feature for program evaluation through its advisory part scores. In the Professional Education Tests there are advisory part scores for: (1) Psychological Foundations of Education, (2) Societal Foundations of Education, and (3) Teaching Principles and Practices. In the General Education Tests there are advisory part scores for: (1) Written English Expression, (2) Social Studies, (3) Literature and Fine Arts, (4) Science, and (5) Mathematics.

It is possible to readminister the Written English, Reading, and Mathematics Proficiency Tests of the California Achievement Tests (CAT) Level 5 to exiting students. Possible statistical analyses include: (1) Pearsonian product-moment correlation coefficient and prediction equations between

BEST COPY AVAILABLE

college grade point average (GPA) and various NTE scores (Ferguson, 1971, pg. 96-119); (2) Pearsonian product-moment correlation coefficient and prediction equations between pre-assessment measures and various NTE scores (Ferguson, 1971, pg. 96-119); and, (3) correlated t-test (Ferguson, 1971, pg. 153-155) or Walsh nonparametric (Siegel, 1956, pg. 83-87) comparison between pre- and post-CAT scores.

Student Guidance

In addition to the semester pre-registration course advisement program, student guidance is to be provided through the Assessment Center primarily by way of diagnostic test evaluation and summary assessment analyses. As indicated under assessment procedures, the Assessment Center is to collect information regarding student progress and provide diagnostic and summary analyses to both the instructor and students. Since each assessment instrument includes a diagnostic feature, the Assessment Center can provide a diagnostic analysis per student in order for the instructor to assign corrective or remedial work. Summary analyses per assessment instrument are to be forwarded to each instructor for within-course student guidance. Furthermore, instructors can adjust their lecture and/or reading material according to pre-test information.

Transfer students are to be evaluated by the Assessment Center with regard to competency in course pre-requisites taken at other institutions.

Program Evaluation and Management

The Assessment Center is to be responsible for providing summary analyses to the Mercy College Education Department and its associated consortium for program evaluation and modification. In addition to Phase I - Pre-Assessment, Phase II - Competency- and Performance-Based Assessment, and Phase III - Post-Assessment, the Assessment Center is to provide summary analyses for: (1) student evaluation of course and competency- and performance-based approach and (2) employer follow-up questionnaire regarding on-the-job performance of graduates.

BEST COPY AVAILABLE

Conclusions

It should be noted that the proposed Assessment Center system and its related statistical procedures have not yet been field tested. Therefore, it is expected that feedback from instructors, students, and consortium members will alter the nature and scope of the assessment instruments and procedures. A fundamental assumption of CBTE programs is continuous system improvement and refinement.

References

Airasian, R.W. Performance-based teacher education: Evaluation issues. In T.E. Andrews (Ed.), Performance education: Assessment. Albany: The University of the State of New York, The State Education Department, Division of Teacher Education and Certification, and Multi-State Consortium on Performance-Based Teacher Education, 1974.

Cornell, F.G. The essentials of educational statistics. New York: John Wiley & Sons, 1956.

Downie, N.M. Fundamentals of measurement. New York: Oxford Univ. Press, 1967.

Downie, N.M. & Heath, R.W. Basic statistical methods. New York: Harper & Row, 1970.

Edwards, A.L. Statistical methods. New York: Holt, Rinehart, & Winston, 1967.

Ferguson, G.A. Statistical analysis in psychology and education. New York: McGraw-Hill, 1971.

Gay, L.R. Evaluation: Data collection, analysis, and reporting. Performance-Based Teacher Education, 1973, 2, 10-16.

Gentry, C., Dunn, T., Myers, D., Cohen, S., Mutterer, M., Yorke, D., Dolinsky, D., & Beckwith, D. For want of an assessment system, CBTE programs are lost. Performance-Based Teacher Education, 1974, 3, 1-11.

Joyce, B.R. Estimating costs of a competency-orientation. In T.E. Andrews (Ed.), Performance education: Assessment. Albany: The University of the State of New York, The State Education Department, Division of Teacher Education and Certification, and Multi-State Consortium on Performance-Based Teacher Education, 1974.

McNemar, Q. Psychological statistics. New York: John Wiley & Sons, 1962.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill, 1956.

Wike, E.L. Data analysis: A statistical primer for psychology students. New York: Aldine-Atherton, 1971.

Copies of this article are available from: Joseph A. Trzasko, Ph.D.; Department of Psychology; Mercy College; Dobbs Ferry, New York 10522.